

خسوسه‌بندی / کامبیز بدیع... [و دیگران]؛ طراح کتاب الکترونیکی
محمود یداللهی خالص، -- تهران: دریاچه، ۱۳۸۴.
۲۷۲ ص.؛ جدول.

ISBN 964-8072-05-1؛ ریال؛ ۲۹۰۰۰

فهرست‌نویسی براساس اطلاعات فیبا.
واژه‌نامه.
کتابنامه.

۱. تجزیه و تحلیل کلاستر. الف. بدیع، کامبیز.

۵۱۹ QA ۲۷۸/غ۹

/۵۳

۱۱۸۸۸-۸۴م

کتابخانه ملی ایران

خوشه بندی

کامبیز بدیع
مریم طایفه محمودی
جوانبخت اسعدی مقدم
لی لی صادقی گل افزانی

ایران، تهران
هزار و سیصد و هشتاد و چهار

انتشارات دریچه با همکاری مرکز تحقیقات مخابرات ایران

تهران، خیابان دانشگاه، شماره ۱۴۶، واحد ۴، صندوق پستی ۱۳۱۴۵-۷۴۹

تلفن / ۶۴۹۱۸۸۵ - ۸۸۸۲۱۲۷۶

Email/Daricheh_pub@hotmail.com



خوشه‌بندی کامپیوتر بدیع

سرگام طایفه محمودی

جوانبخت اسمعی مقدم

لژی صافی گل‌افزار

چاپ اول ۱۳۸۴

حروف نگاری و صفحه‌آرایی
طراح جلد
چاپ جلد
چاپ متن
طراح کتاب الکترونیک

بهار متوجهری
احسان آقامحمد نظری
دیگر
پیام
محمود یداللهی خالص

تعداد
۱۱۰۰ نسخه
۲۹۰۰ تومان



شابک ۹۶۴-۸۰۷۲-۰۵-۱ ISBN 964-8072-05-1

DARICHEH PUBLISHING, Inc.

#4,146 Daneshgah Street, Enghelab Street,

Tehran 13156, IRAN P.O.Box 13145 - 749

Tel/+98 - 21 - 6491885 Fax/+98 - 21 - 6497582

فهرست مطالب

i	پیش گفتار
۱	پیش در آمدی بر خوشه بندی
۷	فصل اول . خوشه بندی شولا
۸	۱.۱ مقدمه ای بر خوشه بندی شولا
۱۰	۲.۱ رویکردی بر خوشه بندی شولا
۱۰	۱.۲.۱ الگوریتم های خوشه بندی شولای کلاسیک.
۱۵	۲.۲.۱ خوشه بندی خطی
۱۶	۳.۲.۱ خوشه بندی صدقی
۲۰	۴.۲.۱ مرزهای اشکال چند ضلعی
۲۲	۵.۲.۱ سایر الگوریتم های موجود در خوشه بندی شولا
۲۹	۳.۱ جمع بندی
۲۹	۱.۳.۱ الگوریتم های خوشه بندی شولای کلاسیک
۲۹	۲.۳.۱ خوشه بندی خطی
۳۰	۳.۳.۱ خوشه بندی صدقی
۳۱	۴.۳.۱ مرزهای اشکال چند ضلعی
۳۹	۴.۱ کار برد خوشه بندی شولا
۴۱	۵.۱ مراجع
۴۹	فصل دوم. خوشه بندی ژنتیکی
۵۰	۱.۲ مقدمه ای بر خوشه بندی ژنتیکی
۵۲	۲.۲ رویکردی بر خوشه بندی ژنتیکی
۵۲	۱.۲.۲ خوشه بندی سخت با استفاده از الگوریتم های ژنتیکی
۵۴	۲.۲.۲ تکه بندی ژنتیکی
۵۷	۳.۲.۲ خوشه بندی ژنتیکی با اشکال هندسی (Genclust)
۵۸	۳.۲ جمع بندی
۵۸	۱.۳.۲ خوشه بندی سخت با استفاده از الگوریتم های ژنتیکی
۶۰	۲.۳.۲ تکه بندی ژنتیکی
۶۱	۳.۳.۲ خوشه بندی ژنتیکی با اشکال هندسی

۶۴	۴.۲ کاربرد خوشه بندی ژنتیکی
۶۸	۵.۲ مراجع

۷۵	فصل سوم. خوشه بندی عصبی
۷۶	۱.۳ مقدمه ای بر خوشه بندی عصبی
۷۸	۲.۳ رویکردی بر خوشه بندی عصبی
۷۹	۱.۲.۳ پیاده سازی الگوریتم BSAS
۸۱	۳.۳ جمع بندی
۸۴	۴.۳ کاربرد خوشه بندی عصبی
۸۸	۵.۳ مراجع

۹۳	فصل چهارم. خوشه بندی وقفی
۹۴	۱.۴ مقدمه ای بر خوشه بندی وقفی
۹۵	۲.۴ رویکردی بر خوشه بندی وقفی
۹۶	۱.۲.۴ الگوریتم های خوشه بندی وقفی مبتنی بر منطق شولا
۹۶	۲.۲.۴ الگوریتم های خوشه بندی وقفی مبتنی بر شبکه عصبی
۹۸	۳.۲.۴ سایر الگوریتم ها
۹۹	۳.۴ جمع بندی
۹۹	۱.۳.۴ الگوریتم های خوشه بندی وقفی مبتنی بر منطق شولا
۹۹	۲.۳.۴ الگوریتم های خوشه بندی وقفی مبتنی بر شبکه عصبی
۱۰۰	۳.۳.۴ سایر الگوریتم ها در ارتباط با خوشه بندی وقفی با تحرک توزیع شده
۱۰۳	۴.۴ کاربرد خوشه بندی وقفی
۱۰۸	۵.۴ مراجع

۱۱۵	فصل پنجم. خوشه بندی خودسازمانده
۱۱۶	۱.۵ مقدمه ای بر خوشه بندی خودسازمانده
۱۱۷	۲.۵ رویکردی بر خوشه بندی خودسازمانده
۱۱۹	۳.۵ جمع بندی
۱۲۴	۴.۵ کاربرد خوشه بندی خودسازمانده
۱۲۶	۵.۵ مراجع

۱۳۹

فصل ششم. خوشه بندی مورد پی

۱۴۰

۱.۶ مقدمه ای بر خوشه بندی مورد پی

۱۴۱

۲.۶ رویکردی بر خوشه بندی مورد پی

۱۴۲

۳.۶ جمع بندی

۱۴۵

۴.۶ کاربرد خوشه بندی مورد پی

۱۴۷

۵.۶ مراجع

۱۴۹

فصل هفتم. خوشه بندی قاعده پی

۱۴۰

۱.۷ مقدمه ای بر خوشه بندی قاعده پی

۱۴۱

۲.۷ رویکردی بر خوشه بندی قاعده پی

۱۴۲

۳.۷ جمع بندی

۱۴۴

۴.۷ کاربرد خوشه بندی قاعده پی

۱۴۶

۵.۷ مراجع

۱۴۹

فصل هشتم. خوشه بندی شاگله پی

۱۵۰

۱.۸ مقدمه ای بر خوشه بندی شاگله پی

۱۵۱

۲.۸ رویکردی بر خوشه بندی شاگله پی

۱۵۲

۳.۸ جمع بندی

۱۵۴

۴.۸ کاربرد خوشه بندی شاگله پی

۱۵۶

۵.۸ مراجع

۱۵۹

فصل نهم. خوشه بندی مدل پی

۱۶۰

۱.۹ مقدمه ای بر خوشه بندی مدل پی

۱۶۱

۲.۹ رویکردی بر خوشه بندی مدل پی

۱۶۳

۳.۹ جمع بندی

۱۶۶

۴.۹ کاربرد خوشه بندی مدل پی

۱۶۸

۵.۹ مراجع

۱۷۱

فصل دهم. خوشه بندی آماری

۱۷۲

۱.۱۰ مقدمه ای بر خوشه بندی آماری

۱۷۳	۲.۱۰ رویکردی بر خوشه بندی آماری
۱۷۴	۳.۱۰ جمع بندی
۱۷۷	۴.۱۰ کاربرد خوشه بندی آماری
۱۸۰	۵.۱۰ مراجع

فصل یازدهم. خوشه بندی یافته ای

۱۸۵	۱.۱۱ مقدمه ای بر خوشه بندی یافته ای
۱۸۶	۲.۱۱ رویکردی بر خوشه بندی یافته ای
۱۸۷	۳.۱۱ جمع بندی
۱۸۸	۴.۱۱ کاربرد خوشه بندی یافته ای
۱۹۰	۵.۱۱ مراجع

پس گفتار

۱۹۷	واژه نامه
-----	-----------

۲۰۹

فهرست جداول

۳۴	جدول ۱-۱ چارچوب مربوط به الگوریتم های خوشه بندی شولا
۴۰	جدول ۲-۱ تعامل میان مفهوم ترکیبی Fuzzy Clustering و موجودیت های ممکنه
۶۲	جدول ۱-۲ چارچوب مربوط به الگوریتم های خوشه بندی ژنتیکی
۶۶	جدول ۲-۲ تعامل میان مفهوم ترکیبی Genetic Clustering و موجودیت های ممکنه
۸۳	جدول ۱-۳ چارچوب مربوط به الگوریتم های خوشه بندی عصبی
۸۷	جدول ۲-۳ تعامل میان مفهوم ترکیبی Neural Clustering و موجودیت های ممکنه
۱۰۱	جدول ۱-۴ چارچوب مربوط به الگوریتم های خوشه بندی وقفی
۱۰۶	جدول ۲-۴ تعامل میان مفهوم ترکیبی Adaptive Clustering و موجودیت های ممکنه
۱۲۲	جدول ۱-۵ چارچوب مربوط به الگوریتم های خوشه بندی خودسازمانده
۱۲۵	جدول ۲-۵ تعامل میان مفهوم ترکیبی Self-Organizing Clustering و موجودیت های ممکنه
۱۳۴	جدول ۱-۶ چارچوب مربوط به الگوریتم های خوشه بندی مورد پی
۱۳۵	جدول ۲-۶ تعامل میان مفهوم ترکیبی Case-Based Clustering و موجودیت های ممکنه
۱۴۳	جدول ۱-۷ چارچوب مربوط به الگوریتم های خوشه بندی قاعده پی

۱۴۵	تعامل میان مفهوم ترکیبی Rule-Based Clustering و موجودیت‌های ممکن	جدول ۷-۲
۱۵۳	چارچوب مربوط به الگوریتم‌های خوشه‌بندی شاکله‌بی	جدول ۸-۱
۱۵۴	تعامل میان مفهوم ترکیبی Schema-Based Clustering و موجودیت‌های ممکن	جدول ۸-۲
۱۶۵	چارچوب مربوط به الگوریتم‌های خوشه‌بندی مدل‌بی	جدول ۹-۱
۱۶۶	تعامل میان مفهوم ترکیبی Model-Based Clustering و موجودیت‌های ممکن	جدول ۹-۲
۱۷۶	چارچوب مربوط به الگوریتم‌های خوشه‌بندی آماری	جدول ۱۰-۱
۱۷۸	تعامل میان مفهوم ترکیبی Statistical Clustering و موجودیت‌های ممکن	جدول ۱۰-۲
۱۸۹	چارچوب مربوط به الگوریتم‌های خوشه‌بندی یافته‌ای	جدول ۱۱-۱
۱۹۱	تعامل میان مفهوم ترکیبی Heuristic Clustering و موجودیت‌های ممکن	جدول ۱۱-۲

فهرست اشکال

۹	شکل ۱ (از فصل اول): هستانگار الگوریتم‌های خوشه‌بندی شولا
۵۲	شکل ۱ (از فصل دوم): هستانگار الگوریتم‌های خوشه‌بندی زنجیری
۵۵	شکل ۲ (از فصل دوم): سه قطعه معرفی شده توسط الگوریتم Sweep
۷۶	شکل ۱ (الف) (از فصل سوم): معماری عصبی
۷۶	شکل ۱ (ب) (از فصل سوم): پیاده‌سازی الگوریتم BSAS
۷۷	شکل ۲ (از فصل سوم): هستانگار الگوریتم‌های خوشه‌بندی عصبی
۹۵	شکل ۱ (از فصل چهارم): هستانگار الگوریتم‌های خوشه‌بندی وافی
۱۱۶	شکل ۱ (از فصل پنجم): هستانگار الگوریتم‌های خوشه‌بندی خودسازمانده
۱۳۰	شکل ۱ (از فصل ششم): هستانگار الگوریتم‌های خوشه‌بندی مورد بی
۱۳۲	شکل ۲ (از فصل ششم): رویکرد مورد بی به خوشه‌بندی
۱۴۱	شکل ۱ (از فصل هفتم): هستانگار الگوریتم‌های خوشه‌بندی قاعده‌بی
۱۵۰	شکل ۱ (از فصل هشتم): هستانگار الگوریتم‌های خوشه‌بندی شاکله‌بی
۱۶۱	شکل ۱ (از فصل نهم): هستانگار الگوریتم‌های خوشه‌بندی مدل‌بی
۱۷۳	شکل ۱ (از فصل دهم): هستانگار الگوریتم‌های خوشه‌بندی آماری
۱۸۶	شکل ۱ (از فصل یازدهم): هستانگار الگوریتم‌های خوشه‌بندی یافته‌ای



پیش گفتار

از سال ها پیش بر این ایده بودیم که ایجاد مهارت در پژوهش، ملزم به ایجاد مجموعه‌ای از ارتباطات خاص در فضای تحقیقات است که در صورت تحقق صحیح، پژوهشگران را به طرز مؤثری، در اهداف ویژه‌ای که در فعالیت های پژوهشی دنبال می کنند، یاری دهد. بخشی از این ارتباطات که مؤسسات معتبر پژوهشی در سطح جهان نیز به آن بها داده اند، نشست های علمی توجیهی و یا جلسات مناظره ای ادواری در حوزه های مرتبط با زمینه کاری مورد نظر است، که بنا به ماهیت پویای خود زمینه ساز ترشح ایده های جدید و پیاده سازی آن به طور مطلوب می باشد. افزون بر این، آنچه هنوز جای فراوان برای کار دارد، استفاده از متون تخصصی ای، متفاوت با متون درسی، مرجع و مقالات تحقیقاتی متداول است، که مفاهیم ریشه ای در یک گستره خاص را، با توجه به انتظارات پژوهشگران، و نهایتاً با هدف تقویت فکری آنان در اختیار بگذارد. شایان ذکر است که آموزش کلاسیک، به رغم آنکه زمینه ساز فضای ذهنی لازم جهت تحلیل ابعاد علمی مختلف در یک حوزه تخصصی است، به تنهایی پاسخگوی پویایی حاکم بر یک فضای پژوهشی نمی باشد-پویایی ای که نشأت گرفته از چندگانگی در مدل ذهنی پژوهشگران، نیازهای پژوهش در حال تغییر و تنوع در طیف مسائل تعریف شده است. این مهم ایجاب می کند که متن ویژه ای با هدف پرورش محققان در زمینه های تخصصی مختلف ایجاد گردد.

به اینگونه بود که بر آن شدیم تا در گروه کاربردهای فناوری اطلاعات (گروه جامعه اطلاعاتی وقت) در پژوهشکده فناوری اطلاعات، برای اولین بار کتابی را با چنین ویژگی هایی در یک زمینه محوری تدوین کنیم. اهمیت وجودی چنین فعالیتی، با توجه به نقش محوری محتوا سازی و تولید متن به عنوان یک فعالیت در حوزه فناوری اطلاعات نیز، قابل توجیه می باشد. این کتاب، با توجه به موضوع محوری

آن که خوشه بندی است، در بسیاری از زمینه های نظری مرتبط با علوم رایانشی که هسته روش شناسی در فناوری اطلاعات را تشکیل می دهند، از قبیل: یادگیری ماشین، بازشناسی الگو، تحلیل سیستم، مدل سازی و شناسایی سیستم ها، ... و در وادی تحقیقات کاربردی نیز در گستره هایی از قبیل متکاوی، زیست مدی (بیومتریक्स)، آموزش هوشمند، محتواسازی برای مقاصد رایا فراگیری، ... کاربری دارد.

از بعد آموزش و پژوهش دانشگاهی نیز، این کتاب ابزار کمک تحقیقاتی مناسبی برای دانشجویان مقاطع کارشناسی ارشد و دکترا در گرایش های مهندسی کنترل، مهندسی مخابرات (سیستم)، مهندسی رایانه، مهندسی سیستم های هوشمند، مهندسی سیستم (به ویژه با تأکید بر گرایش تحقیق در عملیات) و مهندسی پزشکی به شمار می رود. ضمناً، بسیاری از دروس دانشگاهی از قبیل بازشناسی الگو، پردازش تصویر، پردازش سیگنال، مدل سازی و شناسایی سیستم ها، آزمون پذیری و روا داشت خطا، امنیت داده ها، پایگاه داده و بازیابی اطلاعات، یادگیری ماشین، تحقیق در عملیات پیشرفته، تحلیل سیستم های اقتصادی-اجتماعی و... می توانند از این کتاب به عنوان یک مرجع مناسب در مقوله خوشه بندی استفاده نمایند.

یک نکته ممیزه در این کتاب، امکان استفاده از آن بنا به انتظاراتی است که مخاطب، با توجه به مدل ذهنی خود و فعالیت تعریف شده دنبال می کند. به طور مثال، چنانچه هدف از ارجاع به کتاب، آشنایی مقدماتی با مبحث خوشه بندی از یک منظر خاص باشد، کافیت که مخاطب به بخش "مقدمه" از فصل مربوطه مراجعه نماید. هم چنین، برای آن دسته از مخاطبان که در صدد انجام نوعی تحلیل نظری و یا مدل سازی مقایسه ای در ارتباط با مباحث خوشه بندی می باشند، کافیت که بخش "رویکردی بر خوشه بندی" و بخش "جمع بندی" را، در ارتباط با فصولی که در برگیرنده خوشه بندی از منظر های محتمل است، مطالعه کنند.

نکته قابل توجه دیگر اینکه، خوانندگان یک فصل خاص می توانند پیش از ورود به جزئیات، با مطالعه "جدول چارچوب الگوریتم ها" که به طور مجزا برای هر فصل تهیه شده، دید اجمالی ای پیرامون الگوریتم های مورد نظر، امتیازات، نقاط ضعف و گستره کاربرد آن بیابند. به این ترتیب می توان زمان و

انرژی لازم برای فعال ساختن مفاهیمی را که برای مخاطب، با توجه به هدف مورد نظرش در اولویت قرار دارد، کاست. آیین استفاده بهینه از کتاب، به اختصار در جدول زیر شرح داده شده است. به خوانندگان محترم توصیه می شود، در شروع استفاده از کتاب، حتماً به جدول مربوطه رجوع نمایند.

بخش های ۱.۱، ۱.۲، ۱.۳، ۱.۴، ۱.۵، ۱.۶، ۱.۷، ۱.۸، ۱.۹، ۱.۱۰، ۱.۱۱	آشنایی مقدماتی با منظرهای خوشه بندی
بخش های ۲.۱، ۲.۲، ۲.۳، ۲.۴، ۲.۵، ۲.۶، ۲.۷، ۲.۸، ۲.۹، ۲.۱۰، ۲.۱۱، ۳.۱، ۳.۲، ۳.۳، ۳.۴، ۳.۵، ۳.۶، ۳.۷، ۳.۸، ۳.۹، ۳.۱۰، ۳.۱۱	تحلیل نظری و یا مدل سازی مقایسه ای در ارتباط با منظرهای خوشه بندی
بخش های ۴.۱، ۴.۲، ۴.۳، ۴.۴، ۴.۵، ۴.۶، ۴.۷، ۴.۸، ۴.۹، ۴.۱۰، ۴.۱۱	توجیه جنبه های کاربردی در ارتباط با منظرهای خوشه بندی
بخش های ۵.۱، ۵.۲، ۵.۳، ۵.۴، ۵.۵، ۵.۶، ۵.۷، ۵.۸، ۵.۹، ۵.۱۰، ۵.۱۱	انتخاب مراجع مناسب در ارتباط با منظرهای خوشه بندی
پیش در آمدی بر خوشه بندی	آشنایی کلی با مفهوم خوشه بندی و منظرهای گوناگون آن
جدولهای ۱.۱، ۱.۲، ۱.۳، ۱.۴، ۱.۵، ۱.۶، ۱.۷، ۱.۸، ۱.۹، ۱.۱۰، ۱.۱۱	کسب دید اجمالی پیرامون منظرهای خوشه بندی و انواع الگوریتم های مربوطه
جدولهای ۲.۱، ۲.۲، ۲.۳، ۲.۴، ۲.۵، ۲.۶، ۲.۷، ۲.۸، ۲.۹، ۲.۱۰، ۲.۱۱ پس گفتار	تولید ایده ها و موضوعات تحقیقاتی جدید به ازای منظرهای گوناگون خوشه بندی

در پایان، بجاست که از زحمات و تلاشهای فراوان و بیدریغ همکاران اینجانب در تدوین کتاب، خانم مهندس طایفه محمودی، آقای اسعدی مقدم و خانم صادقی گل افزایی صمیمانه سپاسگزاری نمایم. جای شبهه نیست که تحقق این کتاب بدون همدلی و پشتکار این عزیزان میسر نمی گردید. همچنین، از همکاران گروه کاربردهای فناوری اطلاعات، آقای مهندس محمود بدالهی جهت تهیه قالب الکترونیکی کتاب، آقای احسان نظری به خاطر طراحی جلد و خانم بهار منوچهری به واسطه حوصله، دقت و سعه صدر فراوانی که در حروفچینی و صفحه آرایی کتاب به خرج داده اند، و نیز آقایان مهندس نیما ریحانی به

خاطر همفکری های علمی شان در باب خوشه بندی و مهندس محمود خراط به واسطه مشاوره های بیدریغشان، سپاسگزاری نمایم.

در خاتمه ضروری می دانم که مراتب قدردانی خود را از همکاری های بی شائبه و پشتیبانی های معنوی آقای مهندس قنبری (رئیس محترم وقت مرکز)، آقای دکتر صدری (رئیس محترم کنونی مرکز)، آقای دکتر مقدم (معاون محترم پژوهشی وقت مرکز)، آقای دکتر فردیس (معاون محترم پژوهش و فناوری کنونی مرکز) و آقای دکتر آسوشه (معاون محترم توسعه ارتباطات علمی کنونی مرکز) در حمایت از فعالیت تحقیقاتی ای که منجر به شکل گیری این کتاب شده است، اعلام نمایم.

به عنوان نکته آخرین، این کتاب بلاتردید دارای کاستی ها و نقائص فراوانی است که تنها لطف بیدریغ شما خوانندگان محترم در انعکاس آن به نگارندگان، موجب پربارتر شدن چاپ های بعدی آن خواهد گردید. ما را در این رویکرد نو به تدوین متون محقق پرور، که بلاتردید منشأ فواید بسیاری برای جامعه علمی و تحقیقاتی کشور است، یاری دهید.

با آرزوی توفیقات الهی

کامبیز بدیع

مدیر پژوهشکده فناوری اطلاعات

مرکز تحقیقات مخابرات ایران

k_badie@itrc.ac.ir